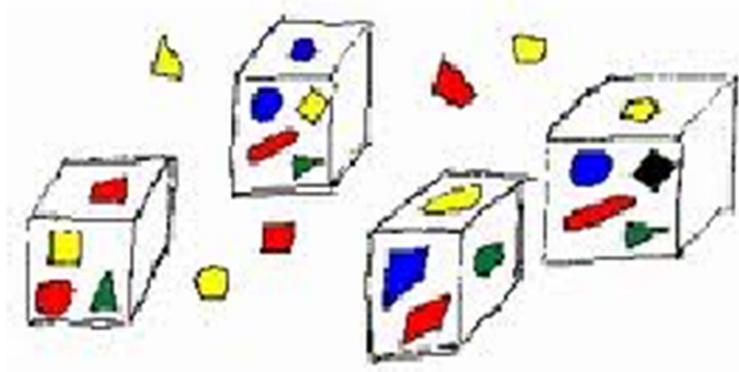


Placement Dynamique de VMs dans le Cloud : Divisez votre Capex par deux !



Makhlouf HADJI

Ingénieur de Recherche - IRT SystemX

ISX-TEO-CCRCCO-0018 (White Paper Cloud Placement)

Table des matières

1	Introduction	4
1.1	Introduction à l'optimisation du Cloud Computing	4
1.2	Objectifs de ce papier	4
2	Définition du problème de placement dynamique de Machines Virtuelles ?	5
2.1	Placement dynamique de VMs dans le Cloud	6
3	Pourquoi doit-on s'intéresser au problème de placement de VMs ?	6
4	Complexité et solutions du placement de VMs	8
4.1	Complexité du problème de placement de VMs	8
4.2	Approches possibles	9
4.2.1	Approches exactes	9
4.2.2	Approches heuristiques	10
4.2.3	Exemples de solutions du marché : VMware	11
4.2.4	Exemples de solutions du marché : Openstack	11
5	Nouvelle Approche	12
5.1	Pourquoi adopter notre nouvelle approche ?	12
5.2	Solution Exacte	13
5.3	Solution Heuristique	14
6	Conclusions et Perspectives	15

Résumé

Ce papier se focalise sur le problème de placement dynamique de machines virtuelles sur les serveurs physiques (*hosts*) d'une infrastructure Cloud. Ce problème consiste à placer les machines virtuelles sur un nombre minimum de serveurs physiques tout en conservant une bonne qualité de service. Après avoir mis l'accent sur les intérêts d'avoir une infrastructure optimisée en termes de réduction du Capex et de l'Opex, on proposera deux solutions efficaces basées sur les plans coupants et sur la théorie des matroïdes, dans le but de résoudre le problème de placement sur des infrastructures de tailles larges. On montrera alors l'efficacité de nos solutions exacte et approchée, permettant d'atteindre des réductions de 50% sur le Capex, et d'environ 40% sur l'Opex. Nos solutions sont totalement compétitives face à celles qui sont proposées sur le marché (VMware) et ou celles qui sont en opensource (Openstack).

1 Introduction

1.1 Introduction à l'optimisation du Cloud Computing

Comme le Cloud Computing devient de plus en plus accessible et populaire, une évolution naturelle consiste à étendre les différents concepts de services, plateformes et infrastructure au provisionnement à la demande afin d'assurer une certaine connectivité entre les ressources virtuelles et les services fournis par le Cloud.

La richesse des offres Cloud ne cessent d'augmenter afin de pouvoir répondre amplement aux différentes demandes des consommateurs. Cette diversification nécessite une bonne urbanisation afin de les présenter sous forme de différents services aux utilisateurs. La multiplication des offres nécessite forcément des efforts de management de l'infrastructure physique afin de pouvoir proposer des politiques de tarifications diversifiées et attractives répondant aux panoplies des demandes.

Aujourd'hui l'offre Infrastructure as a Service (IaaS) est l'une des plus demandées par les utilisateurs de Cloud, et son évolution ne cesse d'augmenter. Cette offre est souvent présentée par des demandes de machines virtuelles (VMs) de différents types.

Pour qu'un fournisseur Cloud de machines virtuelles arrive à proposer une riche tarification de son service, il est nécessaire de réaliser une optimisation plus profonde de son infrastructure physique. En effet, et à titre d'exemple, pour arriver à atteindre les prix ainsi que la richesse des services de AWS (en termes de VMs), il faut passer par une optimisation globale de son infrastructure, et plus loin, réfléchir à des solutions plus réalistes sur des problèmes d'approvisionnement et d'architecture.

Ce papier contribue à cette évolution et à ce besoin de passer par une optimisation de son Cloud pour finalement répercuter les gains obtenus par cette optimisation sur la richesse de l'offre et sur les prix pratiqués. En particulier, on se focalisera sur la partie qui concerne l'optimisation de l'infrastructure afin de réduire la taille et le nombre de serveurs à utiliser tout en assurant une bonne qualité de service aux utilisateurs. Cette problématique relève du problème de placement des VMs sur les serveurs. Par analogie, on peut aussi voir ce problème comme le jeu "Tetris" dont le but est de bien placer les formes qui arrivent, aux bons endroits (voir figure 1).

1.2 Objectifs de ce papier

En plus des facteurs normatifs (émergence de l'internet des services et des standards web), et des facteurs technologiques (virtualisation et réseaux haut débit), les facteurs économiques (crise économique, concurrence, ...) sont derrière l'apparition du Cloud Computing, notamment pour un besoin fort de réduction de coûts associés aux infrastructures des entreprises. Le Cloud Computing est alors une solution adéquate aux différents usages à moindre coût.

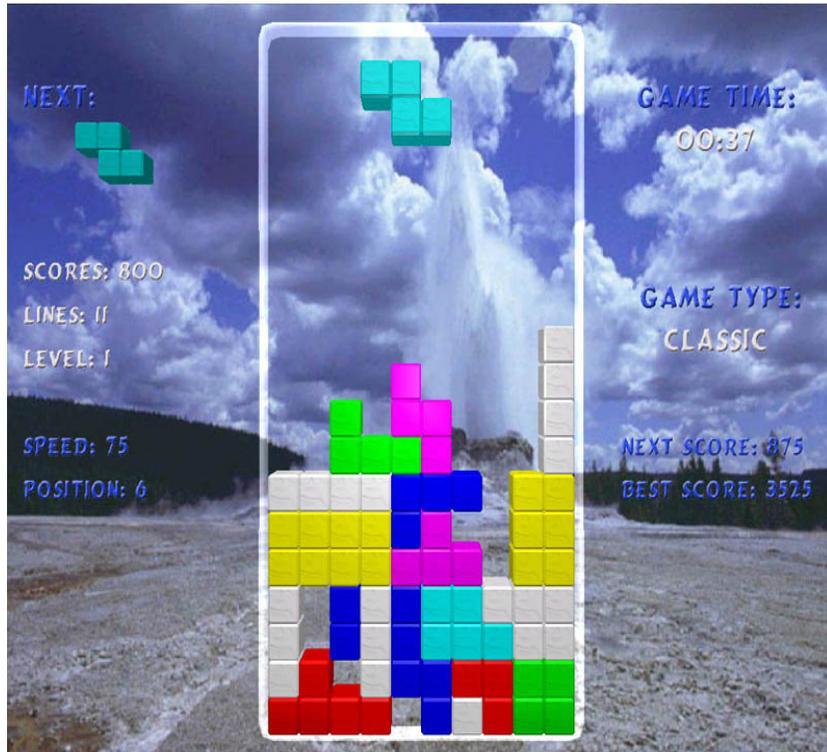


FIGURE 1 – Le jeu Tetris comme analogie au placement

En plus de mettre l'accent sur l'importance et la complexité du problème de placement de machines virtuelles (VMs) sur une infrastructure physique dans le Cloud, ce papier a aussi pour but d'évangéliser des problèmes qu'on rencontrera dans le Cloud, et comment les résoudre efficacement.

Ce papier s'adresse aux fournisseurs de services Cloud Computing, mais aussi à tout ceux qui souhaitent mesurer les performances et les intérêts des solutions d'urbanisation et de management de Cloud Computing passant par des problèmes de placement, et de consolidation des machines virtuelles et des hosts : chefs de projets, DSI, architectes logiciels, développeurs, ...

2 Définition du problème de placement dynamique de Machines Virtuelles ?

Dans cette section, on définit le problème de placement classique, et on donne son équivalente définition dans le Cloud Computing.

2.1 Placement dynamique de VMs dans le Cloud

Le problème de placement est un problème ancien et classique. Il consiste à mettre des objets définis et caractérisés par un volume, dans des boites avec des volumes limités. Le but est de placer tous les objets en utilisant un nombre minimum de boites.

Dans ce qui va suivre, on donnera une analogie entre la définition classique du problème de placement et sa définition dans un contexte de Cloud Computing. En effet, on fera l'analogie en considérant les machines virtuelles comme des objets, et les hosts comme les boites à remplir. Le nouveau problème de placement consiste alors à placer les machines virtuelles dans un minimum de hosts.

Dans les data-centers traditionnels, les applications sont taillées à des serveurs physiques spécifiques qui sont le plus souvent soit sur-chargés ou peu utilisés, et cela dans le but de pallier aux demandes complexes et aléatoires d'une entreprise. Comme conséquence, cette utilisation non rationnelle des ressources induit forcément des coûts opérationnels élevés et des nouveaux investissements en Capex (nombre de serveurs) non pas pour pallier aux pertes de ressources, mais surtout pour pouvoir répondre aux différentes nouvelles demandes et charges. Cette situation dans laquelle on trouve des serveurs non utilisés ou peu utilisés, coûtera alors en termes de mètres carrés occupés, en consommation énergétique, et en toutes les opérations de supervisions et de maintenance associées.

Le Cloud Computing offre aujourd'hui un moyen de découpler les applications du support physique requis pour les faire fonctionner. Cette isolation permet la consolidation des applications multiples en un nombre réduit de serveurs physiques. Ce découplage des ressources est assuré par le concept de machines virtuelles qui encapsule une application avec l'ensemble de ses fonctionnalités.

Les ressources physiques sont mises à disposition aux machines virtuelles par différentes techniques de virtualisation permettant ainsi des opérations de création, destruction et de migration d'une VM. Comme produits de management de VMs, on cite par exemple Xen [10], KVM [1], VMware [2], etc.

Grâce à cette nouvelle capacité offerte par le Cloud et qui consiste à héberger plusieurs applications ou machines virtuelles (VMs) sur le même support physique tout en étant aussi capable de les migrer facilement d'un serveur à un autre, de nouveaux défis sont apparus. Ces derniers incluent la répartition de charge sur l'ensemble des serveurs physiques (PMs), répondant à la question : "*Sur quelles PMs seront placées les VMs ?*". La réponse à cette question relève du problème de placement de VMs tel qu'il est montré sur la figure 2.

3 Pourquoi doit-on s'intéresser au problème de placement de VMs ?

Pendant la période 2007 jusqu'à 2030, la consommation énergétique au niveau international va augmenter de 76% [3], avec une grande contribution en terme de consommation des data-centers.

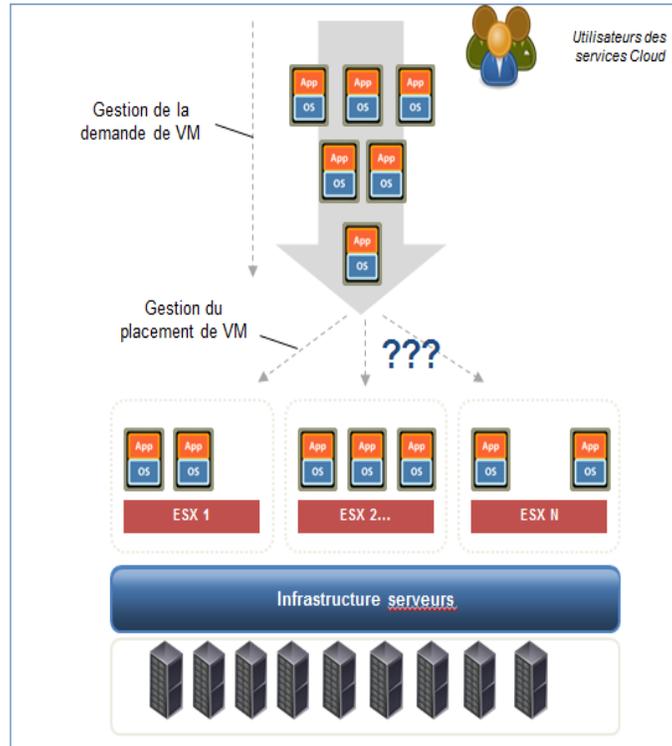


FIGURE 2 – Problème de placement de VMs sur une infrastructure ESX

Ceci induit clairement à réfléchir sur les méthodes à suivre pour réduire cette consommation dans le Cloud. De plus, et selon le rapport Gartner [4], un data-center de taille moyenne consomme en énergie électrique l'équivalent de 25000 foyers. Et selon un rapport McKinsey [5], en 2010, les coûts de consommation énergétique des data-centers sont estimés à 11.5 bn de dollars, sachant aussi que ces coûts doublent chaque 5 ans.

À cette évolution, s'ajoute une information importante sur la consommation d'un serveur en mode veille, et qui est estimée à 70% de sa consommation maximale [6]. Cette perte d'énergie à cause des serveurs en mode veille est considérée comme une cause majeure de la consommation totale d'un data-center. Ceci met clairement le besoin de réduire et de baisser cette consommation par des mécanismes d'optimisation comme celui du placement intelligent des VMs dans le data-center pour solliciter un plus petit nombre possible de serveurs, et d'éteindre le reste.

La figure 3 met clairement en évidence la différence en terme de Capex (nombre de serveurs utilisés) quant à l'adoption d'un placement intelligent de VMs. Cette réduction a des grandes conséquences quant à la consommation énergétique du data-center ainsi qu'aux frais de refroidissement associés.



FIGURE 3 – Exemple de placement de VMs : impacts sur le Capex

4 Complexité et solutions du placement de VMs

Le problème de placement classique est connu comme étant un problème difficile notoire au sens de la théorie de la complexité. Plus de détails sont fournis dans [7]. Dans ce papier, on ne donne pas les explications techniques de cette difficulté mais dans ce qui suit, on explique simplement la combinatoire du placement qui tend vers un nombre exponentiel de cas possibles. Cette multitude de solutions réalisables est alors impossible à explorer en temps raisonnable.

4.1 Complexité du problème de placement de VMs

Pour un petit nombre de VMs et de PMs, il est possible pour un opérateur de réaliser lui même le placement des VMs sur l'infrastructure physique. Dans les cas où le nombre de VMs et de PMs augmentent considérablement, alors la combinatoire qui consiste à trouver les meilleurs emplacements physiques pour les VMs devient trop importante et nécessite forcément une solution algorithmique intelligente et automatique pour atteindre la solution optimale dans les plus brefs délais. En effet, dans le cas de N VMs, on peut facilement atteindre les 2^N cas possibles à énumérer.

Un espace de solutions réalisables de telle dimension reste impossible à exploiter dans des temps raisonnables. C'est pour cette raison, et notamment pour les grandes instances, qu'on va aussi avoir recours à l'utilisation des heuristiques intelligentes dont les solutions sont très proches, voire égales à la solution optimale.

4.2 Approches possibles

Avant de discuter les solutions qui existent sur le marché, on rappelle dans ce qui suit les critères ainsi que les contraintes que l'on souhaite respecter afin de réaliser un placement optimal des workloads. Le placement de VMs peut être décrit brièvement par les points suivants :

- On dispose de M machines physiques, avec des ressources définies en CPU, RAM, Disque par exemple.
- On a des requêtes de workloads de N machines virtuelles à placer. Chaque VM est caractérisée par les mêmes caractéristiques de ressources que les PMs.
- On doit trouver un couplage optimal (ou plus exactement un b -couplage, b étant un nombre entier) entre l'ensemble des VMs et l'ensemble des PMs.

Le but du placement consiste à trouver le couplage optimal entre les VMs et les PMs respectant l'ensemble des ressources disponibles sur le support physique. Plus loin que ça, on peut aussi envisager un placement de VMs en incluant des **contraintes d'affinité** entre des couples de VMs. En d'autres mots, on souhaite trouver un placement optimal respectant le fait que deux VMs par exemple, doivent être placées sur le même serveur physique. Pour des **raisons de sécurité**, on peut aussi avoir des exigences des clients qui consistent à séparer deux VMs sur des serveurs différents. Un autre cas relevant des **contraintes de SLA**, peut exiger de réaliser le placement exclusivement sur un ensemble de PMs bien définies.

Face à un problème d'optimisation combinatoire difficile au sens de la théorie de la complexité algorithmique, il existe deux courants principaux comme solutions : les approches exactes, et les approches heuristiques. Ces deux courants (types d'approches) vont être discutés dans ce qui suit.

4.2.1 Approches exactes

Face à un problème d'optimisation combinatoire difficile, les approches exactes consistent à trouver et à définir d'une manière complète l'enveloppe convexe du problème traité. En effet, dans cette enveloppe, on cherchera la solution optimale en éliminant les parties qui ne contiennent pas notre solution en proposant des plans coupants efficaces (facettes) pour accélérer la convergence vers l'optimum. La figure 4 met en évidence la recherche de la meilleure solution par l'approche exacte.

L'avantage des méthodes exactes est qu'elles renvoient toujours la meilleure solution, mais face à des problèmes dont la combinatoire est très élevée, il conviendrait alors d'éliminer les parties qui ne contiennent pas la solution optimale en proposant de nouvelles facettes au problème. A défaut de proposer une enveloppe convexe complète, on peut dans certains cas ne pas atteindre l'optimum dans des temps raisonnables. C'est dans des cas similaires, qu'on a recours à des solutions heuristiques.

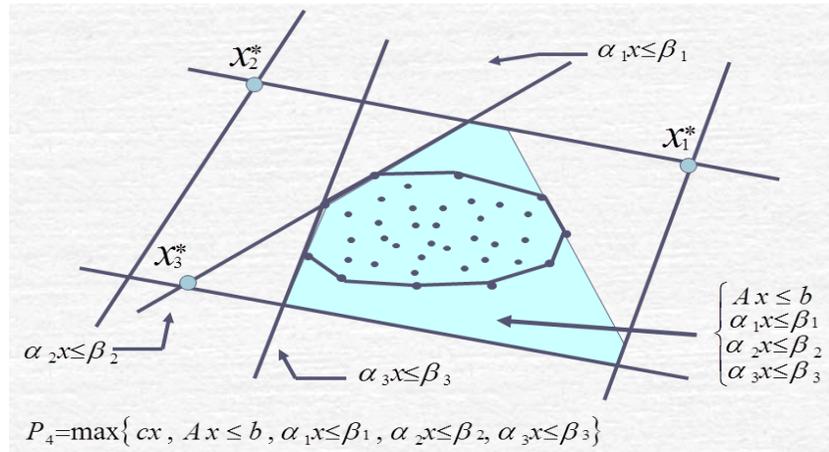


FIGURE 4 – Approches exactes : principe de fonctionnement

4.2.2 Approches heuristiques

Dans les problèmes à complexité exponentielle, il est souvent impossible de réaliser une exploration complète de l'ensemble des solutions réalisables en un temps raisonnable. Une stratégie consiste alors à partir d'une solution initiale réalisable, on propose une amélioration basée sur une modification locale pour qu'elle puisse quitter un optimum local et donc converger le plus vite possible vers l'optimum global. C'est le principe fondamental d'une heuristique.

Concernant le problème de placement dynamique des VMs dans le Cloud, et vu la complexité élevée de ce dernier, il existe alors différentes solutions heuristiques. On en cite quelques unes sans trop rentrer dans les détails. Pour les deux premières solutions, on suppose que les VMs sont triées par ordre décroissant de capacité, puis placées en ordre sur les serveurs :

1. **Best-Fit decreasing** : La VM courante est placée dans le serveur le moins rempli qui peut la contenir.
2. **First-Fit Decreasing** : La VM courante est placée dans le premier serveur qui peut la contenir.
3. **Algorithmes génétiques**.
4. **Méta-heuristiques**, ...

L'ensemble de ces méthodes ont l'avantage principal de converger rapidement vers une solution acceptable, mais qui n'est pas forcément de bonne qualité.

Ce que l'on retient comme avantages et inconvénients des méthodes à utiliser pour résoudre efficacement un problème d'optimisation combinatoire est synthétisé dans la table 1.

TABLE 1 – Table de comparaison

Approches	Approches Exactes	Approches Heuristiques
Avantages	Solution optimale (instances moyennes) Détection des cas faciles	Solutions parfois optimales Mise en oeuvre facile
Inconvénients	Temps de réponse long (cas des grandes instances) Caractérisation difficile	Difficiles à paramétrer

4.2.3 Exemples de solutions du marché : VMware

La solution de VMware est la plus utilisée dans le marché actuellement. Elle est présentée sur la figure 5. Cette solution est constituée de différents modules, et l'intelligence du placement est gérée par le composant DRS (*Distributed Resource Scheduling*). Le DRS est basé sur une solution algorithmique heuristique de type Best-Fit Decreasing qui est décrite auparavant. Au delà du placement, le composant DRS communique avec un monitoring DPM (*Distributed Power Management*) afin d'observer deux critères à savoir la RAM et le CPU. Chaque 5 minutes, une consolidation du système est réalisée afin de pouvoir libérer des emplacements mémoire, ainsi que des serveurs dont le taux d'utilisation du CPU est très faible.

La consolidation proposée par VMware, consiste à migrer à chaud une VM dont les ressources ne sont plus suffisantes, vers un endroit où il y a plus de ressources, et tant qu'il reste encore de la place, on continue à réaliser des migrations des VMs, au point où un effet ping-pong apparaît. Cet effet, consiste à dire qu'une VM peut être migrée d'un serveur X et finir par revenir vers le même serveur. Cette méthode peut causer aussi une perte de données lors de ces migrations, vu le nombre élevé des mouvements qu'on peut constater.

4.2.4 Exemples de solutions du marché : Openstack

Une autre solution qui se veut "*open-source*" et concurrente à la solution VMware, propose une autre algorithmie basée sur des fonctions coûts choisies arbitrairement. Ces fonctions coûts sont attribuées à l'ensemble des serveurs physiques en fonction des ressources disponibles dans chaque host. A titre d'exemple, si un host (serveur physique) dispose de X Mo de RAM, alors une fonction coût peut être simplement $\frac{1}{X}$. Ensuite, on ordonne les hosts du minimum au maximum selon cette fonction coût. Le placement des VMs se fait d'abord dans les hosts de coût minimum et donc de capacité plus large, et on converge au fur et à mesure vers les hosts les plus coûteux ou de capacité restreinte. La figure 6 montre un exemple de solution Openstack. Cette solution est portée par un *scheduler* dit **nova**.

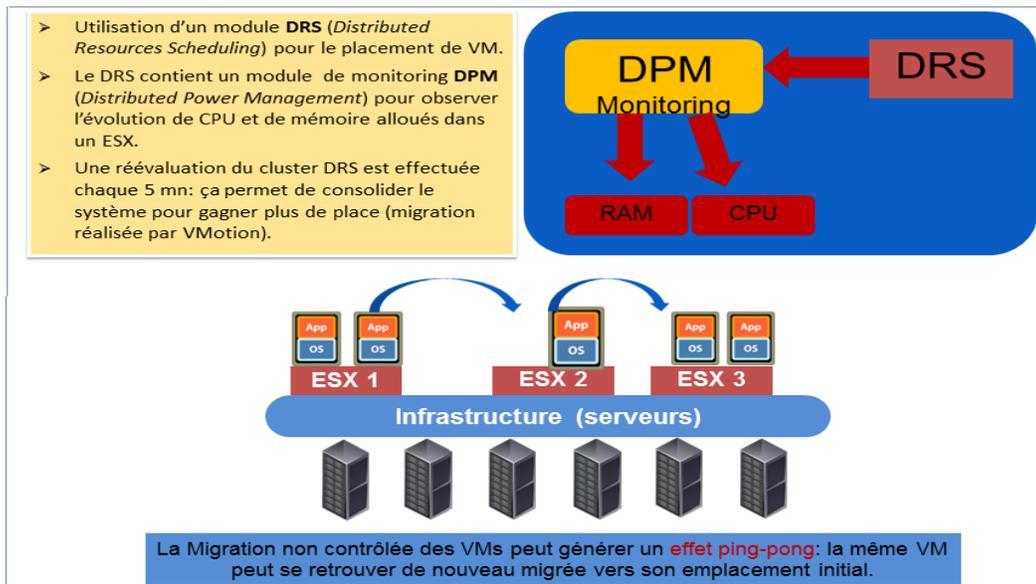


FIGURE 5 – Placement de VMs : solution de VMware

5 Nouvelle Approche

5.1 Pourquoi adopter notre nouvelle approche ?

Nous avons évoqué dans ce papier, l'importance de faire un placement intelligent et dynamique des machines virtuelles. On ne peut plus se contenter des solutions à comportement "Random" ou à des techniques algorithmiques souvent loins de la meilleure solution (optimum). Pour cette raison, on propose dans ce qui suit une nouvelle approche basée sur des méthodes exactes renvoyant tout le temps la solution optimale pour des moyennes instances (environ 2000 VMs), et une solution heuristique basée sur les flots dans les réseaux, qui atteint l'optimum dans les 90% des cas testés.

Avec les solutions qu'on propose ci-dessous, on atteint des gains de près de 50% en Capex, et au moins une réduction de consommation énergétique des serveurs estimée à 40%.

L'adoption de ce type de solutions permet de réaliser des grandes économies sur le Capex ainsi que sur l'Opex. C'est en réalisant ce type d'optimisation, que les grands offreurs du Cloud Computing arrivent à proposer des prix concurrentiels de leurs services (IaaS, PaaS et SaaS).

De plus, les solutions qu'on propose sont bien des solutions Cloud. En d'autres mots, elles sont des solutions qui s'adaptent à la taille des workloads qui oscille vers le haut comme vers le bas (*scalable*).

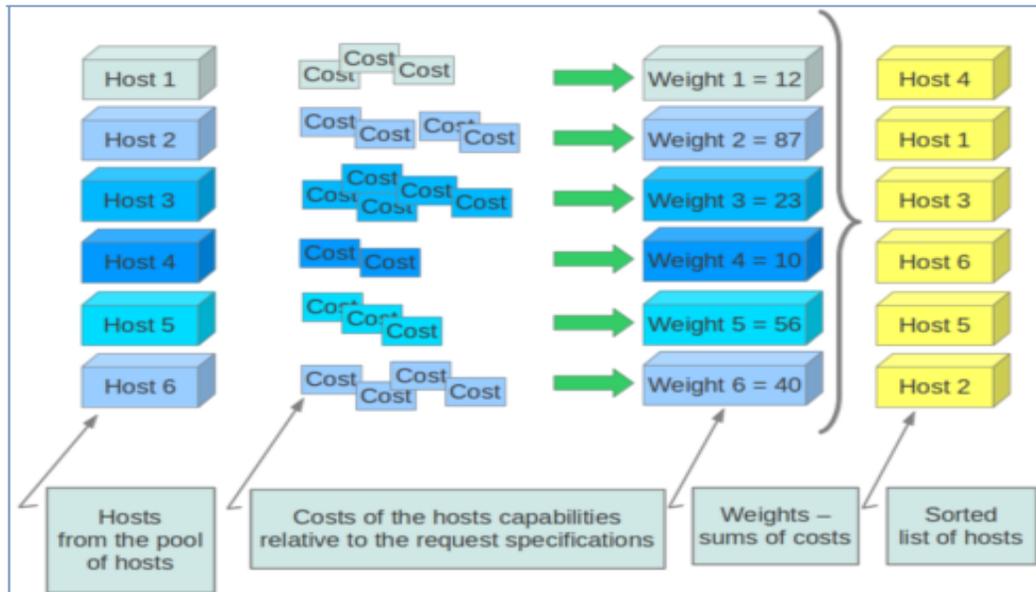


FIGURE 6 – Placement de VMs : solution d'Openstack

5.2 Solution Exacte

Notre approche exacte consiste à définir une enveloppe convexe du problème de placement. On construit alors un programme linéaire sous forme d'un critère à multi dimensions, sujet à différents types de contraintes. Par exemple :

1. **Contraintes de Placement** : Elles représentent les contraintes triviales de placement des VMs tout en respectant la disponibilité des ressources (RAM, Disque, CPU, etc).
2. **Contraintes de Sécurité** : Si pour des contraintes de sécurité, on est amené à séparer deux VMs sur deux serveurs différents, ou deux zones, ou deux data-centers, à la demande de l'utilisateur, alors notre programme linéaire prend ce besoin en compte.
3. **Contraintes SLA** : Ces contraintes sont aussi incluses dans notre modèle exact. Chaque utilisateur peut alors exprimer ses besoins en SLA, et ces derniers seront aussi pris en compte dans notre modèle.
4. **Contraintes de Migration** : Ces contraintes consistent à exprimer des besoins de garder une VM sur le même serveur pour des raisons bien précises.
5. etc ...

La résolution de ce programme linéaire renvoie dans tous les cas, la solution optimale vérifiant l'ensemble des contraintes citées ci-dessus, tout en optimisant le critère pré-défini.

5.3 Solution Heuristique

Lorsque l'on souhaite traiter des instances de grandes tailles, on a recours à une nouvelle heuristique basée sur la théorie des flots.

Dans cette méthode, on construit d'abord un réseau pondéré sur lequel on exprime nos besoins en termes de contraintes et exigences des demandes, ensuite on lance simplement un algorithme de recherche de flot maximum de coût minimum afin de trouver une solution de très bonne qualité.

Cette méthode peut s'avérer efficace car elle nous permet de transformer le problème de placement initial en un problème de flots beaucoup plus facile, et dont une solution souvent proche de l'optimum est caractérisée rapidement.

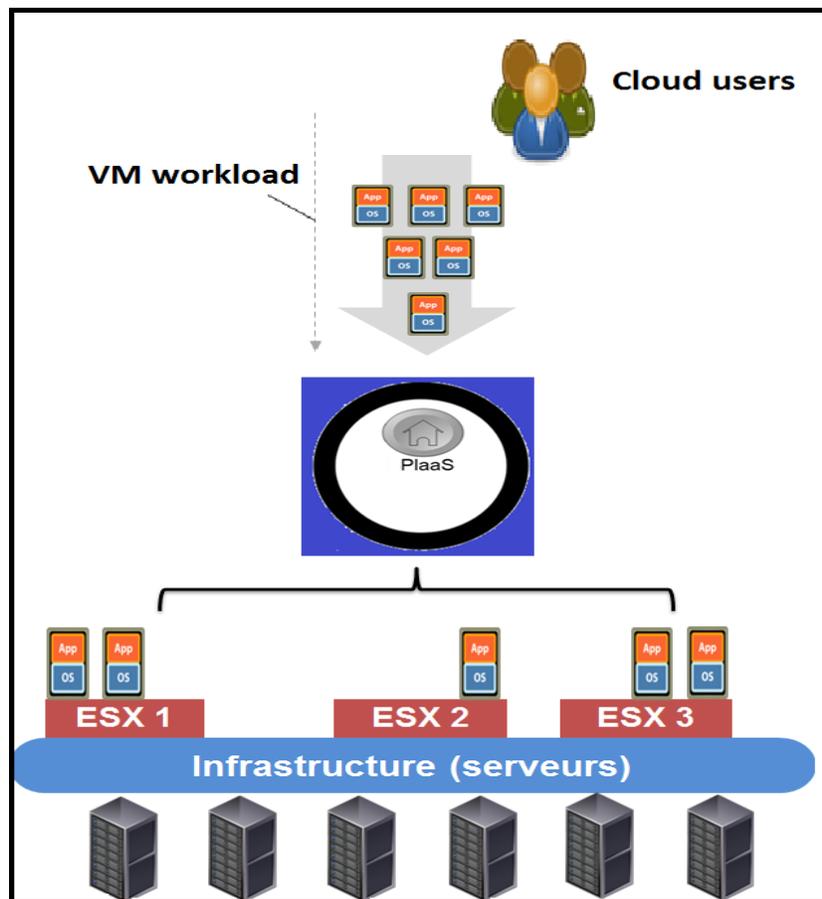


FIGURE 7 – Placement de VMs : nouvelle approche

6 Conclusions et Perspectives

Dans ce papier, on a abordé le problème difficile de placement dynamique de machines virtuelles sur des serveurs physiques. Après avoir mentionné l'importance de réaliser un placement intelligent dans une infrastructure Cloud, nous permettant de faire une optimisation en Capex et en Opex, on a cité différentes solutions existantes sur le marché. On a alors mis l'accent sur les points faibles de ces solutions, et surtout sur les enjeux économiques qui sont derrière un simple placement.

Pour pallier à ce type de défaillances, on a proposé deux approches efficaces : une exacte pour des instances moyennes, et une heuristique avec une bonne qualité de la solution fournie pour de très grandes instances.

Dans un proche temps, on se focalisera sur les solutions du problème de **méta-placement**, qui consiste à réaliser un placement optimal des VMs, avec des contraintes de stockage et de réseau conjointement prises dans la même optimisation.

Il est important de noter que ces problèmes sont encore plus difficiles à cause de la combinaison de différents problèmes initialement complexes. Des approches exactes basées sur la méthodes des plans coupants seront proposées.

Références

- [1] KVM : http://www.linux-kvm.org/page/Main_Page, 2013.
- [2] KVM : <http://www.vmware.com/fr/>, 2013.
- [3] WORLD ENERGY OUTLOOK 2009 FACT SHEET, 2009.
- [4] GARTNER REPORT financial times, 2007.
- [5] J. KAPLAN, W. FORREST, AND N. KINDLER : Revolutionizing data center energy efficiency. McKinsey Company Tech Rep, no. July, p. 15, 2008.
- [6] E. NAONE : Conjuring clouds. Technology Review, vol. 112, no. 4, pp. 54-56, 2009.
- [7] M. R. GAREY AND D. S. JOHNSON : Computers and Intractability : A Guide to the Theory of NP-Completeness. W. H. Freeman 1979, ISBN 0-7167-1044-7.
- [8] NIST, <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf> Novembre 2012.
- [9] M. HOGAN, F. LIU, A. SOKOL AND J. TONG : NIST Cloud Computing Standards Roadmap, Juillet 2011.
- [10] D.E. WILLIAMS AND J. GARCIA : Virtualization with Xen : Including XenEnterprise, XenServer and XenExpress. Amorette Pedersen, 2007.